A Wii-based Gait and Gesture Analysis Project for Teaching Machine Learning

Tom Croonenborghs TOM.CROONENBORGHS@KHK.BE
Biosciences and Technology Department, K.H.Kempen University College, Geel, Belgium

DTAI, Department of Computer Science, KU Leuven, Leuven, Belgium

Peter Karsmakers PETER.KARSMAKERS@KHK.BE
Biosciences and Technology Department, K.H.Kempen University College, Geel, Belgium

SCD/SISTA, Department of Electrical Engineering, KU Leuven, Leuven, Belgium

Abstract

We discuss the student project that accompanies the machine learning course in the Biosciences and Technology Department at the K.H.Kempen University College during the academic year 2011-2012. The goal of this project is to perform gait and gestures analysis on accelerometer data that is recorded by the students themselves using a standard Wii-remote.

1. Introduction

The machine learning course in the Biosciences and Technology Department at the K.H.Kempen University College is given to students following the ICT option in the Master of Science of Applied Engineering: Electronics-ICT. The theory-part of this class consists of (about) twelve 2-hour lectures and uses the handbook by Blockeel (Blockeel, 2010). These lectures give an overview of different machine learning techniques and focus on the main ideas and intuition behind these techniques. This theoretical course is typically perceived by the students as a rather abstract course, especially because these students do not have an extensive mathematical background.

The goal of the lab part of the course is therefore to give the students hand-on experience on real world data that allows them to think about the implications of applying machine learning to real world problems, including the collection of training data, data preprocessing and feature selection. The lab part of the course consists of five 2-hour lab sessions in which the students receive information about data collection and data preprocessing. The evaluation of this practical part of the course is done through the student project. The focus for the students in this project is less on the algorithmic implementation. They receive a collection of tools and procedures and they have to figure out how to combine them in order to develop a successful machine learning application. In our opinion, the students gain a lot of insight during this process.

Section 2 gives an overview of the project while Section 3 provides more details and a description of the steps the students go through when applying machine learning methods to real data.

2. Project Description

In the student project, the students can work in teams of two or three. They have to apply different machine learning techniques to acceleration data acquired by the Nintendo Wii remote. A Matlab toolbox[1] is made available to help with this task. This toolbox contains methods that can be used to record the data, parse the data and perform data recognition. It contains example files for several machine learning methods such as e.g. classification trees (Breiman et al., 1984), kNeirestNeighbours, Dynamic Time Warping (DTW) (Boulgouris et al., 2004) and LS-SVM (Suykens et al., 2002) as well as methods for feature extraction and data transformations.

Appearing in *Proceedings of the ICML12-Workshop "Teaching Machine Learning"*, Edinburgh, Scotland, UK, 2012.

[1]This toolbox is also available at `http://docweb.khk.be/Tom%20Croonenborghs/WiiBasedGaitAndGestureAnalysis.zip`.

2.1. Wii-remote

The Wii remote (or wiimote) is the main controller for the Wii game console by Nintendo. It contains a limited set of motion sensors that allow users to interact with Wii games through movements. The controller contains an accelerometer and optical sensors. For the student project, only the accelerometer data is used. The Wiimote can measure accelerations in three orthogonal directions. It connects via Bluetooth to the Wii console or a PC with appropriate software and sends the measurements, so that these can be analyzed and interpreted, for example, to play a game.

2.2. Gait Analysis

Gait analysis in general is the study of how animals (including humans) walk. The movements and forces that one makes when one walks, crawls etc. can characterize different properties of the animal or person such as e.g. age, weight and general physical condition. Gait analysis is not only used in sports for athletes to improve the efficiency of their technique, but it can be used for all kinds of diagnoses and scientific studies.

The goal of this part of the project is to identify activities on the basis of the acceleration forces that are measured by a standard Wii remote, that is placed in the (right) pocket of the subject's trousers. The different activities are: walking, running and hopping sideways.

2.3. Gesture Recognition

The purpose of this part of the project is to learn a model that can recognize gestures. More specifically, to develop a program that can recognize the digits 0,1,2 or 3 that were recorded by holding the Wiimote more or less horizontally and drawing these numbers in the air.

2.4. Analysis

The students not only have to find those machine learning method that are best suited to solve these problems, but they also have to investigate the necessary conditions to obtain a good classifier such as e.g. the influence of the number of examples, the length of the measurements, the distribution of people in training and test set, etc. The students also have to solve some small additional assignments during the year. An example is an empirical investigation of the selectivity of decision tree induction by adding irrelevant features to the representation of examples.

The students furthermore have to investigate the trade-off between execution speed and classification accuracy. This is especially important for Dynamic Time Warping where the students need to investigate if it is possible to speed up the classification process with a limited accuracy loss by using a smaller training set. Clustering can be applied to find a suitable subset of examples.

2.5. Independent Test Sets

After an initial analysis, the students receive two independent test sets on which they can test the resulting models. These test sets contain some additional challenges, besides the fact that they are recorded by different people. For the gait activity recognition task, other activities are included so that outlier detection can be applied to detect these activities. For the gesture recognition task, the same gestures are recorded as in the student data, but different orientations of the Wiimote are used when recording these gestures. The latter is done to illustrate that application specific data preprocessing and transformations can be useful (as will be shown in Section 3.3.2).

3. Project Implementation

3.1. Data Recording

The first lab session is used to record he data: Each student performed two repetitions of the aforementioned activities and gestures. By doing this, the students not only learn how they can record data themselves but it also motivates them for working on this project.

3.2. Exploration and Visualization

In a first step, the students can familiarize themselves with the complexities of the recorded data and real-life data in general. One way to do this, is to illustrate some of the characteristics of the methods that are discussed in the lectures such as e.g. the decision boundaries of a certain method. To visualize a decision boundary, the recorded data is transformed to a binary feature representation such as e.g. the maximum acceleration in two directions (also see Section 3.3.1).

Consider as an example Figure 1 that shows the decision boundary on the training set created by an QDA classifier for the gestures '0' and '1' using the maximum acceleration in the X- and Y-direction. Figure 2 shows a decision boundary for a kNN-classifier with $k = 3$ applied to the gait dataset using euclidean distance and the median acceleration in the X- and Y-direction as features.

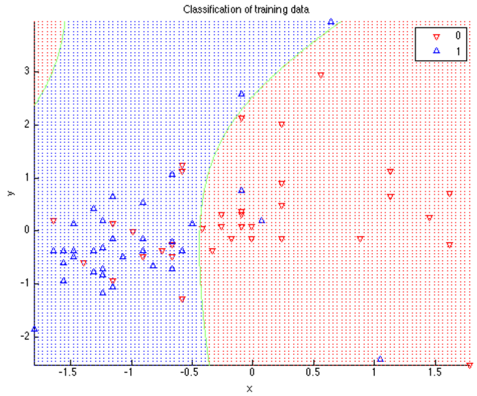


Figure 1. Decision boundary of QDA-classifier using the maximum acceleration in the X- and Y-direction for the gestures '0' and '1'.

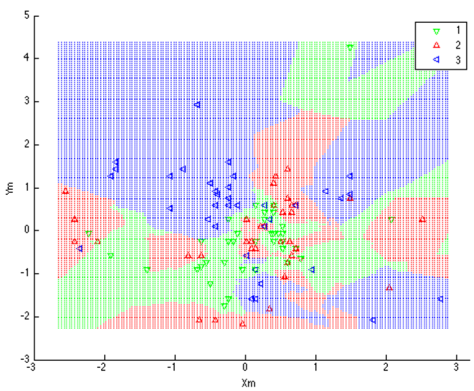


Figure 2. Decision boundary of 3NN-classifier using the median acceleration in the X- and Y-direction for the gait activities.

3.3. Pre-Processing

3.3.1. General Pre-Processing Techniques

Most methods discussed in the lectures operate on an attribute-value representation, therefore the recorded 3-dimensional time-series data needs to be transformed to such a representation by extracting features. The students are given access to simple features such as the minimum, maximum and/or median of each acceleration direction, methods that supply frequency-based features and Linear Prediction Coding (LPC)-coefficients (Makhoul, 1975). Moreover, students are forced to apply and compare proper data standardization techniques on the measurement or data set level. Note that the students are asked to compare these methods using an attribute-value representation with DTW which serves as an example method illustrating classification directly on a time-series representation.

3.3.2. Application Specific Pre-Processing

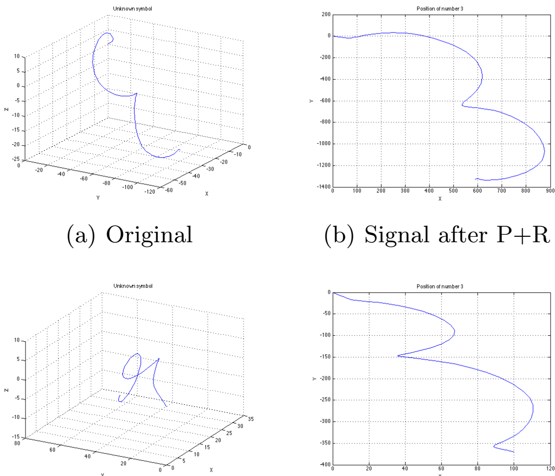
To illustrate that performance can increase when applying application-specific pre-processing, we added some data transformation tools to the toolbox. One of the complexities of accelerometer data is the position and orientation of the Wii-mote during the movements. Note that no gyroscope is used during these measurements. This orientation is especially important for the shorter movements in the gesture recognition task.

It is however possible to try to align all measurements in a pre-processing step. Measuring the decomposition of the G-force over the three orthogonal acceleration directions of the Wii-mote, gives an estimate of the orientation of the Wii-mote. This decomposition can not be measured accurately, but if it is assumed that the Wii-mote is held still in the beginning of the measurement, the (average of the) first (couple of) samples gives a rough estimate. Note that if it is furthermore assumed that all measurements end without acceleration, it is also possible to average over the entire measurement to estimate the decomposition of the G-force.

Besides a rotation in three dimensions, one can apply other transformations as well to align the measurement data. Note that the gestures in principle consist of movements in a two-dimensional plane. The execution of these gestures is however never completely executed in a two-dimensional plane due to user variance and sensor noise. We consider two other transformations that can potentially increase the performance of standard algorithms: 1) 2d-projection to reduce user variance and sensor noise and 2) rotation in this 2d plane to reduce dependence on the specific placement of the accelerometer. The projection operation applies Principal Component Analysis (PCA) and then uses the first two principal components to form the plane in which the gesture was executed. This projection allows for a unique rotation operation that aligns the measured G-force with some reference direction. These transformations have been introduced in (Croonenborghs & Karsmakers, 2011).

To illustrate the effect of these transformations, the students can look at the effect on the computed position information (after integrating the acceleration signal twice). Figure 3 shows the effect of applying the data transformation projection and rotation on two different measurements of the same gesture (symbol '3') but with a different orientation of the

Wii-mote. After applying these transformations, the DTW-distances represent the difference between the examples more accurately.



(a) Original (b) Signal after P+R

(c) Original (different orientation) (d) Signal after P+R

Figure 3. Illustration of the effect of application-specific data transformations.

Other specialized transformations have been included as well such as e.g. the *Time-Delay Embeddings* developed by (Frank et al., 2010) to more accurately estimate gait activities.

3.4. Obtained Results

Some results that illustrate the accuracies that students could obtain using the above mentioned preprocessing and classification methods are given in Table 1 and Table 2. In Table 1 the 10-fold cross-validation (CV) scores are given for both the gait and the gesture data. Note that for the LS-SVM method the RBF kernel is used and a one-vs-one coding scheme is employed to accomplish multi-class classification. Note that method dependent hyper-parameters are automatically tuned for each single CV run.

Method	Acc_{Gait}	$\text{Acc}_{Gesture}$
LS-SVM	97.4(±4.2)%	91.0(±5.6)%
DTW	96.4(±4.6)%	100.0(±0.0)%
kNN	96.5(±6.2)%	87.7(±12)%
Tree	95.7(±4.6)%	67.3(±10)%

Table 1. This table presents the mean and standard deviation of the cross-validation accuracies in terms of correctly classified examples obtained using different classification methods.

Preprocessing	LS-SVM	DTW
no preprocessing	43.8%	50.0%
2-D projection and rotation	56.3%	90.6%

Table 2. This table presents the effect of applying a specific preprocessing on the accuracy in terms of correctly classified examples of an independent gesture test set. Both the persons as the WiiMote orientations differ from the gesture training set used to estimate the classification models.

4. Conclusion

In this paper we gave an overview of the student project that accompanies the machine learning course in the Biosciences and Technology Department at the K.H.Kempen University College. The goal of this project is to perform activity and gesture recognition using accelerometer data recorded with a standard Wii-mote. This project can, in our opinion, motivate students and give them more insight in how to apply machine learning methods to real world problems.

References

Blockeel, Hendrik. *Machine Learning and Inductive Inference*. Acco, 2010. ISBN 9789033482977.

Boulgouris, Nikolaos V, Plataniotis, Konstantinos N, and Hatzinakos, Dimitrios. Gait recognition using dynamic time warping. *IEEE 6th Workshop on Multimedia Signal Processing*, pp. 263–266, 2004.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.

Croonenborghs, Tom and Karsmakers, Peter. Rotation-invatiant detection of 2d-gestures using 3d-accelerometers. In *Proceedings of the Annual IEEE EMBS Benelux Symposium*, pp. 1–2, Leuven, Belgium, December 2011.

Frank, J., Mannor, S., and Precup, D. Activity and gait recognition with time-delay embeddings. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2010)*, 2010.

Makhoul, J. Linear prediction: A tutorial review. *Proceedings of the IEEE*, 63(4):561–580, April 1975. ISSN 0018-9219. doi: 10.1109/PROC.1975.9792.

Suykens, J. A. K., Van Gestel, T., De Brabanter, J., De Moor, B., and Vandewalle, J. *Least Squares Support Vector Machines*. World Scientific, 2002.